

A Prognosis System for Colorectal Cancer

Tiago Oliveira¹, Ernesto Barbosa¹, Sandra Martins², André Goulart³,
João Neves³ and Paulo Novais¹

¹*CCTC/DI, University of Minho, Braga, Portugal*

{toliveira, jneves, pjon}@di.uminho.pt, pg18744@alunos.uminho.pt

²*School of Health Sciences, University of Minho, Braga, Portugal*
sandramartins@ecsaude.uminho.pt

³*General surgery resident at the Hospital of Braga, Portugal*
goulart_andre@hotmail.com, joao.neves@hospitaldebraga.pt

Abstract

The level of uncertainty and incompleteness in the information upon which healthcare professionals have to make judgments has been a subject of discussion in the past, and more nowadays, with the advent of the so-called Clinical Decision Support Systems. This work addresses uncertainty in the postoperative prognosis for colorectal cancer. The interdependence and synergistic effect of different clinical features comes into play when it is necessary to predict how a patient will react to this type of surgery. Using a probabilistic based knowledge representation, a decision support system was conceived in order to provide support for physicians under these circumstances, in particular to surgeons. The solution proposed is based on machine learning on records of cancer patients, incorporating explicit knowledge of experts about the domain. To facilitate access and thus increase its dissemination in the healthcare community, the system is integrated in a wider platform available through a web application.

1. Introduction

Artificial Intelligence in Medicine (AIM) stands for a field in the area of *Artificial Intelligence (AI)* that intends to address problems in health sciences, using *AI* based methodologies and methods for problem solving. One of the subjects reviewed by *AIM* aims at a new picture of uncertainty representation in a clinical context. Prognosis after surgery is very troublesome, mainly due to the biological differences between the individuals, but also due to hidden cause-effect relationships among different clinical features [1]. One of the domains where prognosis is particularly difficult is after ColoRectal Cancer (CRC) surgery. This type of cancer is the third most common worldwide, and is the

fourth cause of death among cancer related illnesses. In Europe [2], it affects predominantly the western countries, a group in which Portugal is included. CRC is the second most common cancer in Portugal [3].

The aim of this work is to present a *Clinical Decision Support System (CDSS)* whose objective is to help healthcare professionals in decision making, during prognosis. In the following section of the paper, the details of CRC and its diagnosis will be described along with the current statistical models in use. Section three provides an insight to the probabilistic model that supports the system and the advantages of its implementation. The functionalities of the system are explained in section four. Finally, the last section of the paper draws some conclusions about the work done so far and points to future directions.

2. Prognosis of Colorectal Cancer

CRC develops in the cells lining the colon when they suffer mutations. In this situation, the mutations cause the uncontrollable growth of these cells and they begin to invade healthy tissues, yielding malignant tumors [4]. They may also spread to other parts of the body by entering the bloodstream or the lymphatic system. The fundamental objectives of cancer prognosis include: the prediction of cancer susceptibility (i.e. risk assessment); the forecast of cancer recurrence and; the expectation of the survival capacity for cancer. Since the scope of this work is focused on the moment immediately after surgery for CRC removal, the outcome used as a reference for assessment will be the after-surgery 30-day mortality. The reason for that is related with the significance it has for the posterior recovery of the patient [5]. Furthermore, if the patient develops any kind of problem during this time period, the responsibility falls upon the surgeon that did the procedure.

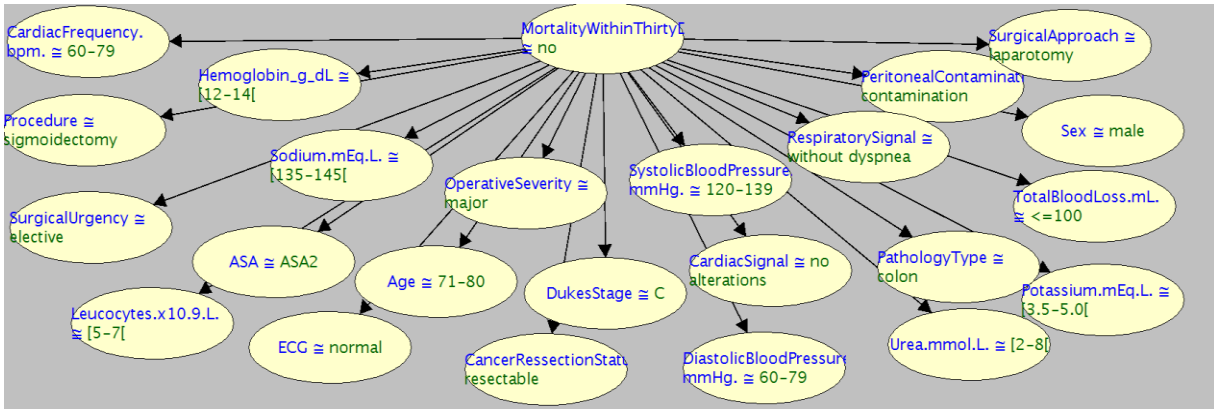


Figure 1. A naïve Bayes model for mortality prediction within 30 days after colorectal cancer resection.

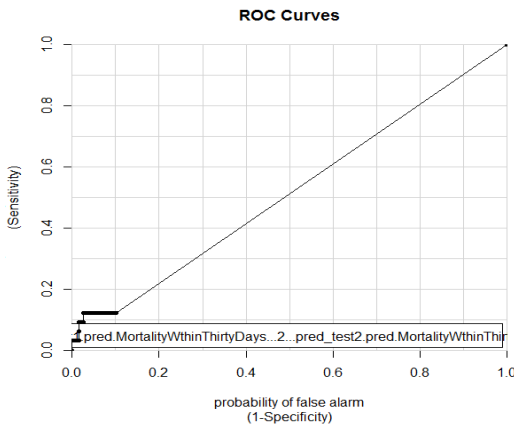


Figure 2. ROC curve for the naïve Bayes classifier.

2.2. Current Models

One of the first scoring systems to forecast surgery was the *Physiological and Operative Severity Score for the Enumeration of Mortality and Morbidity* (POSSUM) [6], which was designed for general surgery. This scoring system is mostly used in the United Kingdom and has a 12-factor, four-grade physiologic score and a 6-factor, four-grade operative severity score. Since the introduction of the original POSSUM system, various modifications have been recommended for the specific requirements of some surgical subspecialties.

Indeed, there is a concern about the applicability of the POSSUM score on different areas of health care. The *Portsmouth POSSUM* (P-POSSUM) [7] system was developed to overcome the problem of over predicting mortality in patients with low risk, using the original POSSUM score. A system for predicting mortality developed specifically for CRC is the *Colorectal POSSUM* (Cr-POSSUM), created in 2004 [6]. This has a better calibration and discrimination than the existing POSSUM and P-POSSUM scores. Within colorectal surgery, oncologic colorectal surgery

is particularly demanding. Patients with CRC are often more susceptible to have problems due to the specific characteristics of CRC, like malnutrition, anemia, and compromised immune systems.

POSSUM, P-POSSUM, and Cr-POSSUM are methods to evaluate the severity of comorbidity (or the appearance of multiple illnesses) and operative factors that might influence surgical outcomes by using complex formulas obtained through logistic regression from observations that included medical comorbidity and severity of operative illness factors. Senagore et al. [6] showed that POSSUM needs to be calibrated in 3 points for each system, and suggest that comparisons between both systems must be done carefully. The Cr-POSSUM score predicts mortality accurately, although missing data from medical records of patients causes variation in the ability to predict the outcome for colon cancer [6][7]. This suggests that the score for an individual patient might be unreliable. Therefore, one must be careful when using scores to forecast individual patient outcomes, since they may influence the selection of practices that may be problematical.

A significant disadvantage of the POSSUM, P-POSSUM and Cr-POSSUM models is that they have not been extensively adopted by the medical community, and therefore, their use is limited. Their performance is poor in populations different from the ones that yielded the sample on which their development was based. This opens the door to the use of AI methods in the elaboration of a general prediction model that may answer the particular needs.

2.1. Significant Indicators

The study of the different views and perspectives about CRC prognosis yielded a set of variables considered to be important for mortality prediction after surgery. The set of variables is divided in two

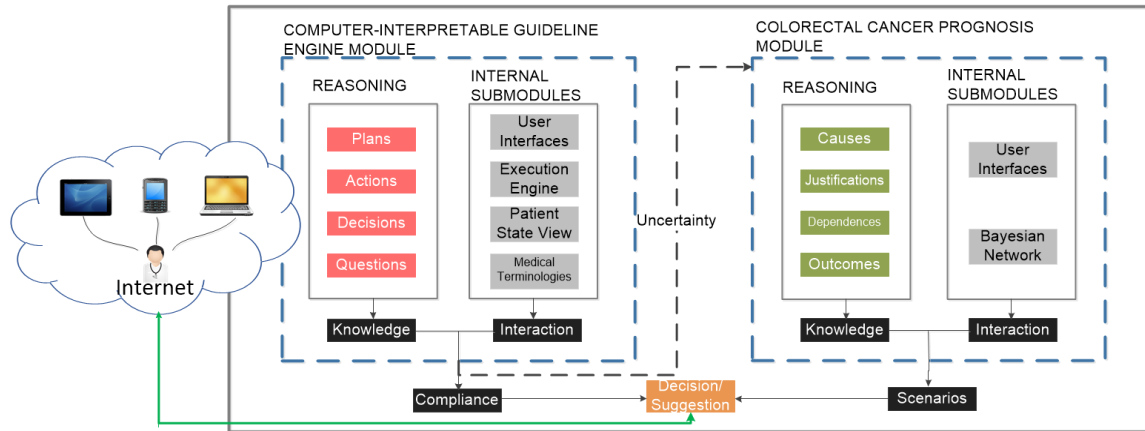


Figure 3. Archetype he archetype for the CDSS with an integration of the CRC prognosis module

main sets: physiological factors and operative severity factors.

The physiological factors are elements that characterize the physical condition of a patient before surgery. This group of factors includes *age*, *sex*, *cardiac signal*, *respiratory signal*, *ElectroCardioGram (ECG) findings*, *systolic blood pressure*, *diastolic blood pressure*, *cardiac frequency*, *levels of substances in the blood* (e.g. hemoglobin, leucocytes, sodium and potassium), and *urea levels*. Also included in the physiological assessment are the *Dukes classification* for cancer and the *American Society of Anesthesiologists (ASA)* physical status classification. One could contend on the use of this cancer classification system over a more recent one, such as the *Tumor Nodes Metastasis (TNM)*, however, given the features available in the data used for examination, the *Dukes classification* [8] had to be considered. On the other hand, the *ASA* score provides a measure of a patient's physical condition, taking into consideration the existence of chronic diseases that may affect his/her quality of life [9].

The operative severity scores reflect the aspects of the surgical procedure that have influence in the patient's recovery. Among those aspects are the *pathology type*, *surgical urgency*, *surgical approach*, *operative severity* (as classified by the surgeon), *total blood loss*, *contamination of the peritoneal cavity*, and the *type of CRC procedure*. It was also added a variable for the *cancer resection status*, i.e., if the surgeon was able remove the tumor completely or not.

These indicators served as a basis for the construction of the prediction models presented in this paper.

3. Probabilistic Model for Prognosis

To predict 30-day mortality, *naïve Bayes* classifiers were considered the best approach. This is a simple

probabilistic model where the evidence variables \mathcal{E} are taken as conditionally independent, given the class variable C .

Using records of 230 patients containing the indicators mentioned above from real patients submitted to *CRC* surgery during 2008 and 2009, in the Hospital of Braga, Braga, Portugal, one constructed two Bayesian classifiers. The models were learned with the tools in the *bnlearn* package of the *R* framework. The first classifier was a simple *naïve Bayes* model (Figure 1), the latter was a *tree augmented naïve Bayes* that differs from the earlier by making an approximation of the dependences among the evidence features and adding directed edges between them.

The results of the *classification error* loss function for the 5-fold cross validation analysis performed on these models are shown in Table 1. Curiously, the tree augmented model that takes into consideration the dependences between the input variables has a worst performance, presenting higher values for the minimum, mean and maximum classification error. The classification errors of the naïve Bayes model were relatively low. However, the *receiver operating characteristic* (ROC) curve for the naïve Bayes model, presented in Figure 2, shows that for higher values of sensibility (or recall) the model loses specificity. As the capability to detect true positives increases so does the probability of issuing a false alarm. This happens to the point where it may be considered that the model is operating at random (given the *sensibility*=*1-specificity* curve). This could be explained by the small size of the dataset and the high number of variables being considered. Moreover, a death is a rare event so its statistical significance in the current dataset is reduced. The data collection is still being processed in order to gather more accounts for further improvement of the model.

4. Clinical Decision Support for CRC

The foremost objective of *CDSSs* is to aid healthcare professionals in the critical moments of the clinical process analysis, providing them with different alternatives concerning the best way to manage diseases or its treatment procedures. When developing such a system it is important to devise how it will be made available to practitioners, in order to reach its target population and how the value of existing models may be integrated into it. Figure 3 shows the archetype of a *CDSS* where the developed model for *CRC* prognosis is integrated.

Table 1. Loss function for 5-fold cross validation.

Classification error	Model	
	naïve Bayes	tree augmented naïve Bayes
Min	0.00000	0.2826
Mean	0.08261	0.3565
Max	0.15220	0.4783

The *CDSS* planned in Figure 3 consists of modules, one for *Computer-Interpretable Guidelines (CIGs)* and another for *CRC* prognosis. The reasoning component of the *CIG* module uses knowledge kept in different logical theories or constructs that, together, build a *Clinical Practice Guideline (CPG)*. This knowledge is used in its inner modules to provide suggestions according to the information fed by the user. However, there are situations that *CPGs* are unable to foresee. As it was previously discussed, the 30 day prognosis after *CRC* surgery is not an easy task, so *CPGs* cannot accurately deal with this situation. There is a need for a tool that is more dynamic and interactive in the exploration of knowledge. In this sense, the *BN* model presented above may offer an interesting complement to the recommendations of *CPGs*.

5. Conclusions and Future Work

This work shows that the combined use of different knowledge representation formalisms may help to capture more accurately the universe of discourse and the relationships among its actors. The classifiers showed here may offer a complement to rule-based systems such as *CIG* execution engines. However, the model is still under development and needs further refinement. It is necessary to re-evaluate the *naïve Bayes* and the tree augmented *naïve Bayes* with the introduction of new variables (and the removal of not so relevant ones) such as treatments performed before the procedure (e.g., chemotherapy, radiotherapy) and other that recently have been equated for *CRC* prognosis, such as the years of experience of the

surgeon. Short term goals also include the addition of a morbidity variable representing diseases a patient is likely to develop after the surgery. This variable will certainly be influenced by the other issues discussed in this paper and, in turn, influence the prediction of mortality within 30 days after surgery. Indeed, the tree augmented naïve Bayes seems (again) a suitable choice.

9. Acknowledgements

This work is partly funded by national funds through the FCT – Fundação para a Ciência e a Tecnologia (Portuguese Foundation for Science and Technology) within project PEst-OE/EEI/UI0752/2011". The work of Tiago Oliveira is supported by a doctoral grant by FCT (SFRH/BD/85291/2012).

10. References

- [1] P. J. F. Lucas, H. Boot, and B. Taal, "A decision-theoretic network approach to treatment management and prognosis," *Knowledge-Based Systems*, vol. 11, no. 5–6, pp. 321–330, 1998.
- [2] J. Ferlay, P. Autier, M. Boniol, M. Heanue, M. Colombet, and P. Boyle, "Estimates of the cancer incidence and mortality in Europe in 2006," *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO*, vol. 18, no. 3, pp. 581–92, Mar. 2007.
- [3] C. G. Pinto, A. T. Paquete, and I. Pissarra, "Colorectal cancer in Portugal," *The European journal of health economics : HEPAC : health economics in prevention and care*, vol. 10 Suppl 1, pp. S65–73, Jan. 2010.
- [4] S. Winawer, R. Fletcher, D. Rex, J. Bond, R. Burt, J. Ferrucci, T. Ganiats, T. Levin, S. Woolf, D. Johnson, L. Kirk, S. Litin, and C. Simmang, "Colorectal cancer screening and surveillance: Clinical guidelines and rationale—Update based on new evidence," *Gastroenterology*, vol. 124, no. 2, pp. 544–560, 2003.
- [5] M. Castro and S. F. Martins, "Evaluation of quality parameters of rectal cancer surgery at the Coloproctology Unit of Hospital de Braga," *Journal of Coloproctology (Rio de Janeiro)*, vol. 31, pp. 362–371, 2011.
- [6] A. J. Senagore, A. J. Warmuth, C. P. Delaney, P. P. Tekkis, and V. W. Fazio, "POSSUM, p-POSSUM, and Cr-POSSUM: implementation issues in a United States health care system for prediction of outcome for colon cancer resection," *Diseases of the colon and rectum*, vol. 47, no. 9, pp. 1435–41, Sep. 2004.
- [7] M. Horzic, M. Kopljar, K. Cupurdija, D. V. Bielen, D. Vergles, and Z. Lackovic, "Comparison of P-POSSUM and Cr-POSSUM scores in patients undergoing colorectal cancer resection," *Archives of surgery (Chicago, Ill. : 1960)*, vol. 142, no. 11, pp. 1043–8, Nov. 2007.
- [8] Y. Adachi, M. Mori, Y. Maehara, and K. Sugimachi, "Dukes's classification: a valid prognostic indicator for gastric cancer," *Gut*, vol. 35, no. 10, pp. 1368–1371, 1994.
- [9] "ASA Physical Status Classification System." [Online]. Available: <http://www.asahq.org/ForMembers/ClinicalInformation/ASA-Physical-Status-Classification-System.aspx>. [Accessed: 17-Feb-2013].